Open source tools in support of a multi-center prospective biomarker study in prostate cancer

Robert A. Dennis, Ph.D., Rey Encarnacion, Andrew Helsley, Sung Hong, Khy Huang, Alan G. Robinson, M.D. UCLA InfoShare, David Geffen School of Medicine, Los Angeles, CA

Abstract. The open source software tools developed at UCLA are the informatics infrastructure supporting translational research efforts of the 11 National Cancer Institutes’ Prostate SPORE sites. The Codebook is a caBIG compliant application that manages PHI (e.g. names, medical record numbers, etc.) separately from research data. Research data are collected using a flexible clinical trials management suite called pTracker. Our poster will present how these two applications are implemented and function in this prospective multi-center study.

Introduction. The Inter-SPORE Prostate Biomarker Study (IPBS) is a multi-year, multi-institutional study that serves as a demonstration project for National Cancer Institute’s (NCI) new paradigm for biomarker discovery and translation, namely, prospective collection of biospecimens from a defined, diverse population, using standardized protocols, and careful annotation of sources of pre-analytic variation. The purpose of the IPBS is to investigate new potential biomarker alternatives to PSA. A total of 700 pretreatment men with prostate cancer will participate and will provide biological samples (tissue cores taken at the time of needle biopsy and blood samples). Participants consent to three years of annual follow-up.

Informatics. A key objective of the IPBS is to develop and pilot an informatics infrastructure to annotate and integrate specimen banks to enhance the quality and availability of specimens and associated data for the broader scientific community. A research coordinator at each Specialized Programs of Research Excellence (SPORE) site is using a local instance of a Cancer Bioinformatics Grid (caBIG) Java application called CodeBook* to add patients to the study. The Codebook at each SPORE site sits behind that site’s firewall. When a patient is added to a Codebook, a message is sent to a central, shared instance of pTracker requesting that a research subject record be created. No PHI other than dates of procedures is entered into pTracker. pTracker is an open source clinical trials management software toolkit developed at UCLA*. IPBS pTracker development efforts began with a requirements document, specific use cases, and printed versions of the case report forms. Two separate workflows were defined to manage both the accrual of patients into the study and the tracking of batches of biomaterials periodically sent from each of the SPORE sites to a central pathology lab and then to the individual biomarker labs for analysis. The combination of CodeBook and pTracker is in use currently, and a long-term goal of this effort is to share biomaterial across the SPORE sites using the federal caBIG model (see figure 1). Each site will install an instance of caTissue and populate it with data exported from pTracker (represented in red in figure 1). CaTissue will allow for distributed queries across the federated sites.

Discussion. All 11 NCI Prostate SPOREs are collaborating on this study and the establishment of an informatics infrastructure to support additional translational research efforts. Establishing an infrastructure that will facilitate efforts like the IPBS is an important informatics goal.

* UCLA InfoShare is funded by the National Library of Medicine IAIMS Operations Grant G08 LM007851 and has supported the development of pTracker and Codebook.